Deep Statistical Learning for Climatology

Guoqi Qian

School of Mathematics and Statistics
The University of Melbourne, Australia

Joint work with Jasper Wijnands, Jackson Kwok and Yuri Kuleshov

Climatology and Deep Statistical Learning

- Climate is a spatial-temporal distribution of weather.
- Climatology studies the trends and variability in climate.
- Stochastic and statistical techniques are indispensable tools.

Two research tasks

§1 Modelling, identification and forecasting in tropical cyclone

§2 Antarctic temperature trend and variability profiling

Motivation of Task 1

- The Australian Bureau of Meteorology (BoM) issues operational tropical cyclone (TC) seasonal forecasts for the Australian region (AR) and the South Pacific Ocean (SPO).
- Calling for better forecasting methods due to improved understanding of TC system.

Our aims in Task 1

Aim 1. Determine the drivers and indices of TC genesis by statistical variable/model selection methods at a range of TC informed spatio-temporal resolutions

Aim 2. Develop multilevel stochastic climate models to delineate important spatio-temporal attributes of TC

Aim 3. Develop coherent statistical modelling procedures for accurate prediction of TC activities at seasonal or shorter timescales in the Southern Hemisphere.

Work completed

1. Some key drivers of TC genesis were identified through *variable selection methods* in graphical model structure learning (*Monthly Weather Review, 2016*).
2. A *spatial-temporal cubic spline model* was developed to represent the near-surface wind speed field in TCs (*Applied Mathematical Modelling, 2016*).
3. A TC seasonal forecasting procedure based on *support vector regression*ecast was developed (*Mathematics of Climate and Weather Forecasting*, 2015).

Key variables driving tropical cloud cluster (TCC) into TC

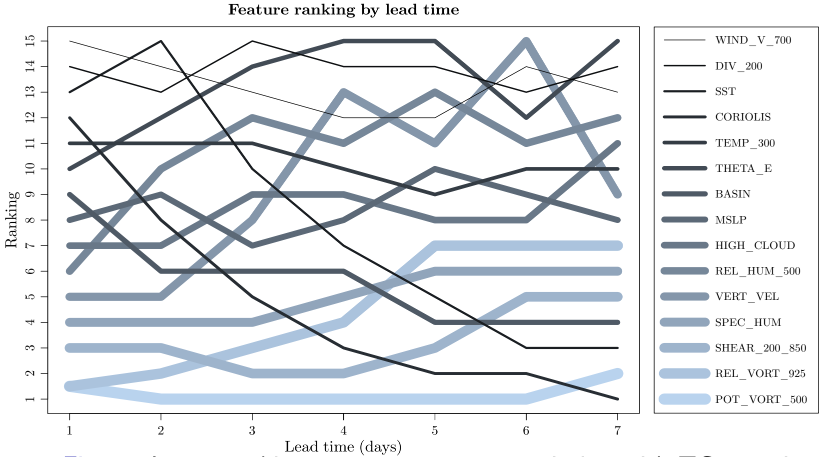


Figure: Lower ranking means stronger association with TC genesis

Description of some potential drivers

Name	description
BASIN	basin
CORIOLIS	Coriolis parameter
DIV_200	divergence (measured at 200 hPa)
HIGH_CLOUD	high cloud cover
MSLP	mean sea level pressure
POT_VORT_500	potential vorticity (measured at 500 hPa)
REL_HUM_500	relative humidity (measured at 500 hPa)
REL_VORT_925	relative vorticity (measured at 925 hPa)
SHEAR_200_850	wind vector shear (measured between 200 and 850 hPa)
SPEC_HUM	specific humidity (averaged over the selected pressure levels
SST	sea-surface temperature
TEMP_300	air temperature (measured at 300 hPa)
THETA_E	equivalent potential temperature difference between surface
VERT_VEL	vertical velocity (averaged over the selected pressure levels)
WIND_V_700	meridional component of wind (measured at 700 hPa)

Summary of the highest ranking features

Rank	1 day	2 days	3 days	4 days	5 days	6
1	REL_VORT_925	POT_VORT_500	POT_VORT_500	POT_VORT_500	POT_VORT_500	POT_V
2	POT_VORT_500	REL_VORT_925	SHEAR_200_850	SHEAR_200_850	CORIOLIS	COR
3	SHEAR_200_850	SHEAR_200_850	REL_VORT_925	CORIOLIS	SHEAR_200_850	S
4	SPEC_HUM	SPEC_HUM	SPEC_HUM	REL_VORT_925	BASIN	BA
5	VERT_VEL	VERT_VEL	CORIOLIS	SPEC_HUM	SST	SHEAR
6	REL_HUM_500	BASIN	BASIN	BASIN	SPEC_HUM	SPEC

TC Hamish, category 5, northeast Cairns, 03/2009

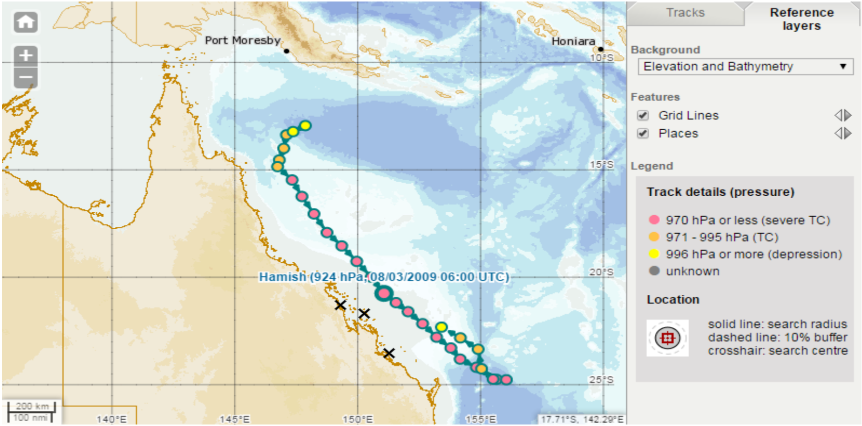


Figure: Best track data for TC Hamish

Cubic spline surface model of wind speed field

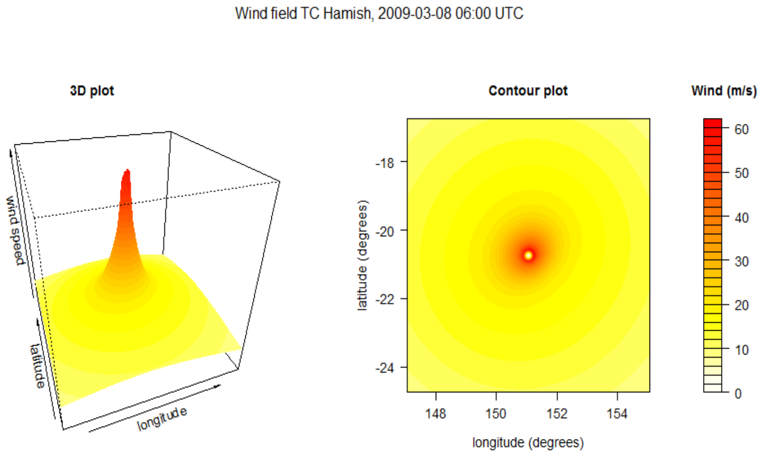
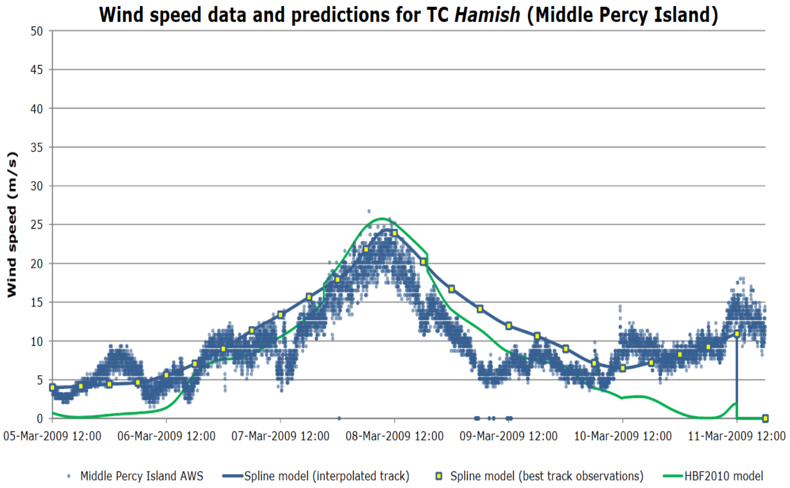
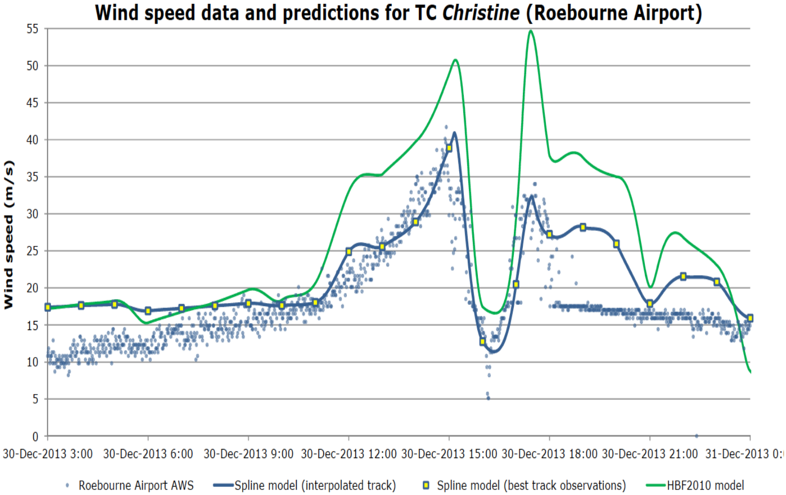


Figure: TC Hamish wind field modelled by cubic spline surface

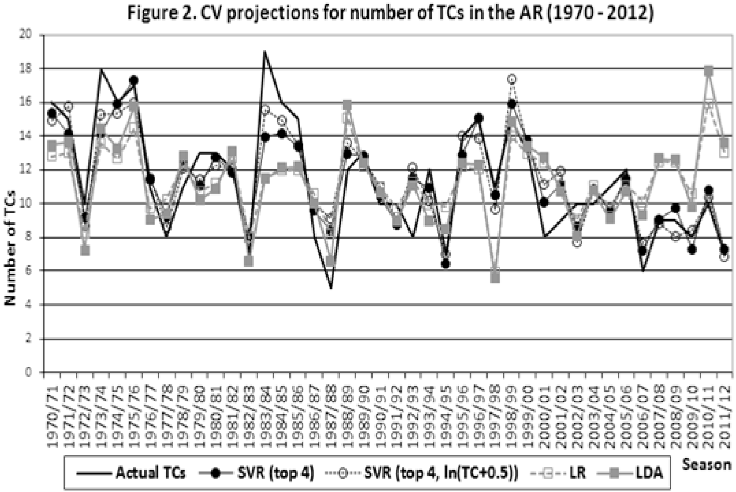
Validating the cubic spline surface model



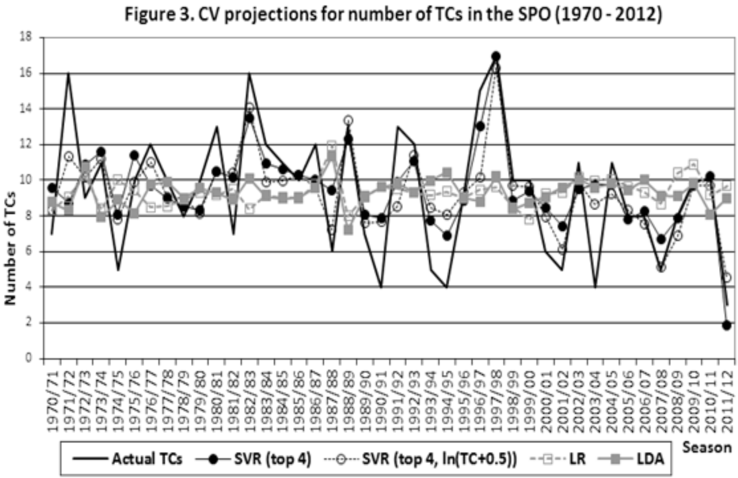
Validating the cubic spline surface model



Cross-validated forecasts of TC for the Australian region



Cross-validated forecasts of TC for South Pacific Ocean

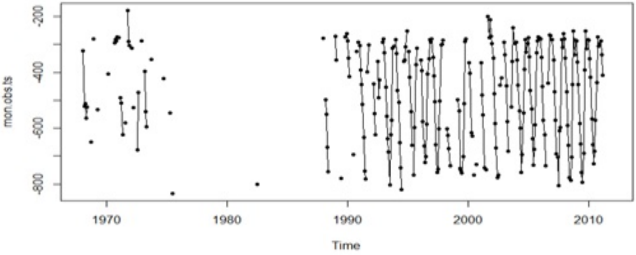


Temperature trends in free atmosphere difficult to analyse

- Historical radiosonde data on temperature in free atmosphere over Antarctic available at `www.bom.gov.au/ant/`.
- Collected at 22 stations, 16 altitude levels, over 60 years.
- Profiling the long-term temperature trends to provide an accurate assessment of the observed climate change.
- A difficult task due to inhomogeneities, sparseness and missingness in the data.

A typical dataset: Casey89611_1000_night.dat

- Only 2890 records over 44 years (=16071 days).
- Hence 13265 days (i.e. 82.54%) had no observations.
- The observations were also sparsely distributed.



Profile of temperature trends at 9 atmospheric pressure levels derived from the radiosonde data from 9 stations(preliminary results)